Assessment of the Weight of Evidence of Formaldehyde as a Human Carcinogen

Prepared by
Lorenz R. Rhomberg, Ph.D.
Gradient
20 University Road
Cambridge, MA 02138
617-395-5000
lrhomberg@gradientcorp.com

On behalf of the
Formaldehyde Council, Inc.
1300 Wilson Blvd.
Arlington, VA 22209

and

Arclin
5865 McLaughlin Rd
Mississauga, Ontario L5R 1B8
Canada

October 16, 2009

Table of Contents

Introduction

Formaldehyde (CAS No. 50-00-0) is listed in the *Eleventh Report on Carcinogens* (ROC) as "reasonably anticipated to be a human carcinogen" based on the NTP's judgment of limited evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in laboratory animals (NTP, 2005). Formaldehyde was nominated by NIEHS for possible reclassification in the 12th ROC as an agent known to be a human carcinogen. A *Draft Report on Carcinogens Background Document for Formaldehyde*, dated September 3, 2009 (NTP, 2009) has been produced to review the evidence, and public comment on the Background Report and the reclassification issue has been invited (74 *FR* 44845, August 31, 2009).

Gradient, an environmental and risk assessment consulting firm, is pleased to respond to this invitation with the following comments. The effort to produce these comments was supported by the Formaldehyde Council, Inc. and Arclin, but the comments are my own.

The chief impetus for a possible reclassification of formaldehyde's status is the report of apparent increases in lymphohematopoeitic cancers in a National Cancer Institute (NCI) study of workers in ten formaldehyde-producing or formaldehyde-using industries (Hauptmann *et al.*, 20032004), and more specifically on a recently updated analysis of the NCI cohort based on 5 further years of follow-up (Beane Freeman *et al.* 2009).

Consideration of a possible reclassification of formaldehyde raises several key questions:

1. Is there a new reason (not evident before the listing in the 11th ROC) that the human data on *other*, non-lymphohematopoietic cancers – principally nasopharyngeal cancers – now support a "known" classification when they previously had been judged to constitute "limited evidence in humans"?
2. Do the reported effects on lymphohematopoietic cancers in the NCI study – when considered in terms of epidemiologic analysis – constitute a robust finding of an increase? That is, are the results clear, consistent, and attributable to formaldehyde exposure rather than to chance statistical anomalies, confounders or co-exposures?
3. Is it biologically plausible that inhaled formaldehyde should be able to cause lymphohematopoietic cancers, in view of the apparent inability of inhaled formaldehyde to pass into the systemic circulation and reach potential target tissues? That is, is there a biologically plausible mechanism by which formaldehyde could reach and transform hematopoietic stem cells? If such a mechanism were operating, what other observable consequences should it have for effects of formaldehyde inhalation in humans and animals, and are these other effects indeed observed?

4. Why is it that inhaled formaldehyde shows no sign of affecting hematopoietic cancer risk in any rodent study, despite lifetime high exposure (and despite the appearance of other tumor responses at the high doses)?

5. In view of the above, do the NCI lymphohematopoietic cancer results constitute or not constitute addition to the weight of evidence for potential human carcinogenicity?

6. In view of a weighing and synthetic interpretation of all of the evidence – that existing previously and that which has been adduced since the 11th ROC; the various human study results and the evident lack of robustness of their findings; the biologically implausible modes of action that must be hypothesized; and the array of negative animal bioassay results – is a reclassification of formaldehyde's human carcinogenicity warranted? In particular, can a conclusion of "known human carcinogen" be sustained?

To address such overarching, weight-of-evidence questions requires delving into the particulars and scientific details of all the areas of evidence. Doing so in depth is beyond the scope of my present comments, but other commenter's examine the component pieces of evidence and their interpretation in detail. My aim is to refer to those detailed, in-depth analyses as they bear on how the various parts should be melded into an overall weight-of-evidence judgment. (The numbering of sections below refers back to the six numbered questions above).

1. The Human Data on Nasopharyngeal Cancers Constitute Limited Evidence

Nasopharyngeal cancers are rare in humans, and so the numbers of observed and expected cancers are low, making study-by-study or plant-by-plant findings susceptible to chance effects or skewing by co-exposures to agents other than formaldehyde. If, following a method that has been employed by ATSDR, one adds up the observed cases and the expected cases across studies, the numbers are nearly identical , showing that collectively the studies do not point to an excess of nasopharyngeal cancers.

Marsh and Youk (2005) and Marsh *et al.* (2006) have shown the sensitivity of results to the particular statistical model and dose measure. Moreover, meta-analyses of nasopharyngeal cancers clearly show that the apparently positive results all stem from the inclusion of data from one plant ("Plant 1", the Wallingford plant) with a locally anomalous result (Marsh *et al.* 2005; Bachand *et al.* 2009). If Plant 1 is excluded, the results from the remaining nine plants show no significant effect of formaldehyde on nasopharyngeal cancer risk, a result confirmed by meta-regression analysis (Bachand *et al.* 2009).

Moreover, the workers in the studied industries have significant exposures to other agents that are plausible causes of nasopharyngeal cancer. Many of the studied nasopharyngeal cancer cases were

among workers in the metalworking industry who were exposed to metal dust, mineral acid, and sulfuric acid fumes in addition to formaldehyde. Marsh *et al.* (2007) conducted a nested case-control study that found it likely that the excess mortality of nasopharyngeal cancer in the Plant 1 cohort was not due to formaldehyde but rather to these other metalworking exposures.

In sum, the human evidence on nasopharyngeal cancers does not constitute a compelling case for an effect of formaldehyde. It is not only possible but seems likely that apparent effects are better attributable to the anomalous Plant 1 result, a likely function of confounding with metal dust and acid exposures among workers with previous employment in metalworking industries.

Nasopharyngeal cancer would have little impact on the weight of evidence analysis were it not for the parallel observation of nasal tumors in rats after chronic inhalation of high concentrations of formaldehyde (Kerns *et al.*, 1983; Monticello *et al.* 1996), which now serve as the sole basis for cancer risk assessment. This is because the animal data demonstrate the biological plausibility of similar tumors in humans. The question is how much this apparent parallelism should contribute to the weight accorded human studies, particularly now that more data are available on the likely mode of action. Extensive study of the rat tumor responses has shown that cancers are only elevated at high exposure levels and in conjunction with marked cytotoxicity and histopathology changes in the affected tissues (Monticello *et al.* 1990, McGregor *et al.* 2006) and at the human exposure levels involved, it is highly unlikely that a process similar to that in rats would operate. As shown in Figure 1, these tissue changes, and the nasopharyngeal tumors, appear only at ongoing exposures of 6 ppm or higher, and human sensory irritation (which occurs at about 1 ppm) would limit human exposures to much lower levels on an ongoing basis.

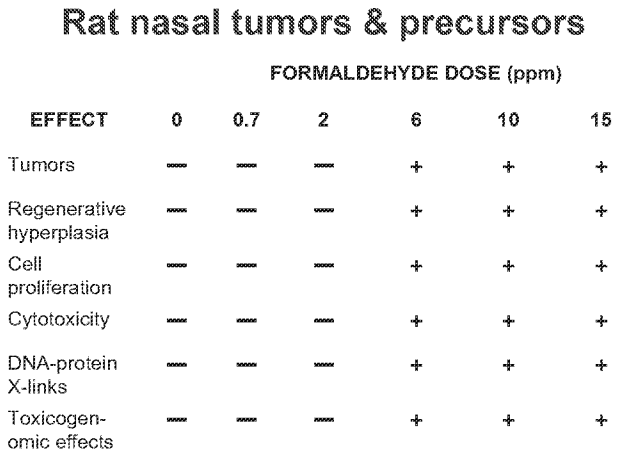

Figure 1

To conclude, the principal difference between now and the 11th ROC in the evidence regarding nasopharyngeal cancers is that meta-analysis, plant-by-plant study, and evaluation of non-formaldehyde exposures makes it increasingly likely that any apparent elevation is not attributable to formaldehyde. That is, the ability to exclude chance, bias or confounding factors has only increased, and the human evidence from nasopharyngeal cancers is still, at best, to be judged as "limited" under the NTP criteria.

2. The Lymphohematopoietic Cancers in the NCI Study do not Constitute a Robust Finding

As noted earlier, the chief difference in the available data on formaldehyde and cancer between the 11th ROC judgment and the present decision is the advent of further follow-up on the NCI cohort regarding lymphohematopoietic cancers (Beane Freeman *et al.*, 2009). The main focus among these has been on leukemias, and my comments will focus there, though similar comments may be made about other specific cancers. Beane Freeman *et al.* report a statistically significant increasing trend in leukemias with higher peak formaldehyde exposures, and a significant difference between those workers with the highest peak exposures (≥ 4ppm) versus the lowest category. No association was found when cumulative or average exposure were used as exposure metrics. Beane Freeman *et al.* represents the latest update on a cohort that has been published on previously with some inconsistent signs of leukemia effects, and the results just noted are what remain upon the most complete follow-up.

There are other epidemiologic studies that have examined hematopoietic cancers, and some have reported indications of elevations, but these studies have been available since before the 11th ROC and do not show

a clear or consistent pattern of effects. The main new information on these other studies consists of reanalysis and their incorporation into meta-analyses that (as will be discussed further below) tend to show that there is no tendency or indication of elevation of any particular hematopoietic cancer across the available studies.

The existing animal bioassay data show no sign of an effect on hematopoietic cancer risk. Thus, any use of such cancers in the weight of evidence must depend on human data alone. Moreover, as discussed more fully below, any conclusion that formaldehyde is elevating human leukemias needs to be reconciled with the lack of such response in animals exposed to much higher and long-lasting doses, despite the broadly similar machinery of hematopoiesis across mammalian species.

Thus, the question of whether formaldehyde should be reclassified as a known human carcinogen largely hinges on the interpretation of hematopoietic cancers, and leukemia in particular – as reflected in the body of epidemiologic literature and most especially as reflected in the sole new results, the update of the NCI study.

For a number of reasons, examination of the evidence on these questions suggests that the putative effect of formaldehyde on leukemias is not a robust and meaningful finding. These are set out and discussed in the following.

2a. The Focus on "Peak" Exposures in the NCI Cohort is Arbitrary, Post Hoc, and Inconsistent with Needed Mode-of-Action Hypotheses

In the NCI cohort, no effect on leukemia incidence was found for formaldehyde exposure calculated as estimated cumulative exposure or average exposure. It is not clear how such measures of total exposure could fail to show an effect if there is indeed an impact of formaldehyde; even if higher exposure intensities are of more consequence, those experiencing such higher air concentrations with any repeatability would have higher cumulative and average exposure, and hence these measures ought to show an association as well.

Moreover, any dependency for risk on high peaks (over 4ppm) needs to be consistent with mode-of-action hypotheses. As discussed further below, in order to explain how formaldehyde, which does not get beyond the immediate respiratory tract tissues, can affect hematopoiesis, it has been proposed that formaldehyde's genotoxicity affects mobile hematopoietic stem cells while they are in the respiratory tract

or that inhaled formaldehyde can somehow be complexed and delivered systemically to hematopoietic tissues (which will lead to its dilution and diffusing of peaks). It is not clear how such processes, without which formaldehyde's leukemogenic action is biologically implausible, would be related to peak exposure but unaffected by ongoing lower exposures.

When the same set of data is analyzed in multiple parallel ways using different models, groupings, or summary measures, the meaning of statistical tests becomes distorted by the multiple-comparisons problem. That is, if enough alternatives are tried, some might be "significant" by chance alone (since, at a criterion of $p = 0.05$, even when there is no effect, 5% of comparisons are ruled "significant"). The NIOSH data were analyzed many different parallel ways (average or cumulative or peak exposure; pairwise comparisons or trends; using internal or external controls; with or without a "nonexposed" group in the trend test; individual tumor types or various measures of combined tumors, *etc.*). Unless a correction for multiple comparisons is made, finding marginal significance in one or a few such comparisons is not surprising even when there is no true effect. There was no *a priori* reason to focus on peak exposures, and so the result for peaks can at most be a hypothesis-generating observation to be tested on future data. Otherwise, it is *post hoc* and arbitrary.

2b. "Peaks" were not Measured in the NCI Study

The NCI study did not actually measure peak exposures; the inference that some workers were exposed to peaks was based on job descriptions and "the likelihood that a high-exposure task or event would occur." There is no way to know whether workers placed in different categories really experienced the peaks inferred for them.

2c. The Total Number of Formaldehyde "Peaks" was not Associated with Leukemia Risk

In the NCI analysis, a person was classified by the highest ever (inferred but not measured) peak exposure, but no distinction was made between whether a working life had one peak or repeated peaks. Another separate analysis *did* examine leukemia risk as a function of the cumulative number of high peaks inferred as having been experienced (*i.e.*, by considering duration in job categories inferred to have high peak levels), and this analysis showed no effect of formaldehyde on leukemia. If peak exposure is really biologically important and the explanation of the reported effect with any peaks, then more peaks should have had more effect than fewer, but this is not the case.

2d. The Analysis of "Peaks" was Only Significant when a "Non-Exposed" Category with Lower-than-Background Leukemia Risks is Added

The NCI analysis has a "low" (>0 to < 2.0 ppm) a "medium" (2.0 to < 4.0) and a "high" (≥ 4.0) category for the inferred lifetime peak level, and analyses were conducted on risk vis-à-vis the low level as a standard. These analyses showed no significant effect. If, however, a further "zero" category is added, comprising workers from the facilities that were presumably unexposed, the trend for leukemia *vs.* "peak" became significant, as did the contrast between the "high" *vs.* the "zero" (but not *vs.* the "low") group.

Even though the "low" group included people down to zero as the lifetime "peak" exposure, the leukemia risk for the "zero" group was markedly lower. People classified as "zero" must have systematically different job descriptions than those in the "low" category (for which peak exposure could be as low as zero and still admit them into the "low" group), and so the comparability of these groups is in question. Moreover, the "zero" group appears to have leukemia risks that are notably smaller than the general population. Indeed, when analyses were done on an SMR basis (comparing risks to outside referent populations rather than using the "low" group as a standard), there was no significant effect.

Similar arguments can be cited for other lymphohematopoietic responses in the NCI study. In short, it appears that the reported significant relation of "peak" formaldehyde exposure and leukemia risk depends entirely on a lower-than-usual leukemia rate in the "zero" group rather than to any effects among exposed people. Since "peaks" were inferred possibilities rather than actual exposures, since the inferred number of high peaks was not related to leukemia, since the cumulative or average exposure was not related to leukemia, and since the finding of a significant effect relies on picking apparently positive results out of a sea of similar analyses showing no effect, one must conclude that the report of a relation between formaldehyde exposure and leukemia in the NCI cohort is not a substantive or compelling finding.

This is so even limiting the consideration to its merits as a study of empirical associations of formaldehyde and hematopoietic cancers. In addition, as discussed below, there are the issues of lack of corroborating or consistent effects in the larger body of human studies, the lack of parallel effects in animal studies, and the biological plausibility issue regarding whether a mode of action can be imagined by which formaldehyde in the respiratory tract can affect hematopoietic tissues.

The foregoing discussion focused on the NCI study, but the body of epidemiologic studies as a whole tends to show no collective pattern, as discussed in the following.

2e. In the Major Human Studies of Leukemia, Observed Deaths Match Expected

As shown in Figure 2, if one adds up the observed leukemia deaths in the three major epidemiologic studies of formaldehyde exposure and also adds up the expected cases for those cohorts, the numbers agree very closely. This methodology was deemed superior to assessing weight of the evidence by ATSDR (1999).

Formaldehyde & leukemia

COHORT	# WORKERS	O	E
NCI	25,000	116	~116
Coggon	14,000	12	13.2
Pinkerton	11,000	24	~24
TOTAL	50,000	152	153.2

Figure 2

That is, in the three major studies, 152 cases were observed and, based on demographics in these cohorts, 153.2 would be expected. This simple analysis does not substitute for a meta-analysis, but it does show that among some 50,000 formaldehyde industry workers examined, there is no evident sign of a marked excess of leukemias. Of course, some of these workers had lower exposures and some higher, but any tendency for the higher-exposed workers to have excess leukemias must come at the expense of a deficit of expected cases among lower-exposed workers. That is, the totals constitute a "zero-sum game" in that any excess over expected cases in one setting must be balanced by a deficit elsewhere.

2f. Meta-Analyses Show No Consistent or Collective Effect on Leukemia Across Studies

A number of meta-analyses have been done on the body of epidemiologic studies concerning formaldehyde and leukemia. Only the most recent of these, Bachand *et al.* (2009), includes the most recent update to the NCI study. This comprehensive and thorough study took measures to examine and avoid selection bias, and it ran extensive sensitivity analyses on its key analytic choices.

Bachand *et al.* found no evidence of significant heterogeneity among studies. For cohort studies, REs ranged from 0.43 to 1.60 for leukemias, with all but one study having 95% confidence intervals including 1.0. For two case-control studies the RE was 0.98 (95% CI: 0.70, 1.36) for Blair *et al.* (2001) and 1.40 (CI: 0.25, 7.91) for Partanen (1993). Meta-regression showed the overall leukemia RE was 1.05 (95% CI: 0.93, 1.20). No effect was seen, even in the higher exposure studies.

According to Bachand *et al.* (2009), earlier meta-analyses took inadequate notice of the potential for heterogeneity. Some may have had issues with selection bias. The results of the three meta-analyses of leukemias and formaldehyde appearing since 2004 can be briefly summarized as in Figure 3.

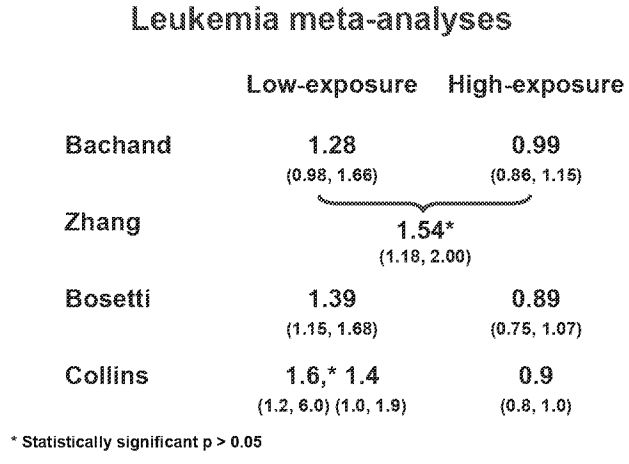

Figure 3

Of these, all but Zhang *et al.,* (2008) allow separate consideration of low-exposure and high-exposure industries. It is notable that the high-exposure industries have, if anything, a lower collective indication of effect than the low-exposure industries. Collins and Lineker (2004) found a small but significant effect among embalmers (1.6, CI:1.2,6.0) and a marginally significant effect among pathologists and anatomists

(1.4, CI:1.0,1.9), both low-exposure professions, but no effect among higher-exposed industrial cohorts. Moreover, these medicine-associated job categories may be affected by diagnostic bias. Bosetti *et al.* (2008) found a similar pattern although no effect was statistically significant.

Zhang *et al.*, (2008) found a significant effect across industries, but they had an unusual means of selecting and combining studies: they used different measures of exposure, selecting only one from each study even if several were examined, resulting in their selection of peak exposure for some studies, average exposure for others, cumulative exposure for still others, and exposure duration for the balance. Moreover, if several categories or levels of exposure were examined, they took data from only the highest among them, and what constituted a "high" category also varied considerably among studies, depending on how each study established gradations of exposure. As a consequence, the comparisons across studies are very heterogeneous, and it is not clear whether a comparable question was being examined in each case. These results should be interpreted with caution, especially in view of their lack of concordance with other meta-analyses.

3. It is Biologically Implausible that Formaldehyde Would Cause Lymphohematopoietic Cancers

A key member of Bradford Hill's criteria for making judgments about causality is biological plausibility (Bradford Hill, 1965). If a biologically mechanism by which the agent could cause the effect cannot be identified or even plausibly hypothesized, it is difficult to justify any conclusion other than that apparent associations of an agent and a disease are actually due to other factors, confounding, or chance. This is especially so when (as has been argued for formaldehyde and lymphohematopoietic cancers above) the observed association itself is tenuous, inconsistent, and susceptible to identified plausible alternative explanations.

It is important to be clear that the argument is *not* that a particular mechanism of carcinogenic action need be positively identified before an agent can be considered a carcinogen. Rather, it is to point out that at least plausible hypothetical mechanisms are implicitly but necessarily assumed whenever the associations observed in epidemiologic studies are taken to be evidence of a causal effect of the agent on the disease state.

That is, any judgment in favor of causality must include the scientific judgment that a material means for the causal process exists. When there is significant evidence *against* the existence of *any* plausible causal

process – as I would argue is true in the present case – any weight-of-evidence evaluation that favors causality must find that the observed associations are so strong and compelling, and that alternative explanations of their appearance are so implausible, that they overcome and refute the apparent impossibility of a mechanism to connect the exposure to the disease.

It is widely understood that, because formaldehyde is so reactive, inhaled vapors cannot get past the immediate respiratory tissues at the site of uptake due to rapid detoxification. That is, there is no systemic distribution of the inhaled gas. This is verified by the fact that even extensive and substantial animal and human inhalation exposures are unable to cause an increase in measured blood levels (Heck *et al.*, 1985; Casanova *et al.*, 1988). No evidence has been presented that formaldehyde gets to sites outside of the respiratory tract. This is evidence against the notion that, by some yet unidentified means, inhaled formaldehyde can be complexed or protected from reaction, distributed to other tissues, and then locally reactivated to cause genotoxic or other interactions in distant hematopoietic tissues.

All established leukemogens act on the bone marrow, the primary site of hematopoiesis. Even for agents with genotoxic action, they all lead to readily observed marrow toxicity. The lack of a plausible means for formaldehyde to reach marrow or any other systemic tissue proposed for involvement in leukemogenesis would rule out formaldehyde as an agent capable of causing leukemia. This argument has been compelling during past reviews of formaldehyde's potential human carcinogenicity, and in these reviews, owing to the biological implausibility, human studies of leukemias and other hematopoietic cancers have not been considered even limited evidence for human carcinogenicity (IARC, 2006, NICNAS, 2006).

Zhang *et al.* (2008) have hypothesized three means by which they propose that formaldehyde could attack hematopoietic stem cells: (1) inhaled formaldehyde gas is converted to hydrated methanediol which enters the systemic circulation and reconverts to reactive formaldehyde in the marrow; (2) a small proportion of stem- and progenitor cells arising in the marrow enter the systemic circulation and are exposed and affected by formaldehyde as the blood passes through the respiratory tissues, with damaged cells later being reincorporated into the marrow; and (3) pleuripotent stem cells in the nasal or oral cavity linings could be affected by formaldehyde and then migrate to marrow, where they enter the hematpoietic process.

These hypothesized mechanisms have been critically reviewed and refuted as plausible possibilities by Golden *et al.* (2006), Pyatt *et al.* (2008) and Goldstein (2009). I believe that other submitted comments examine the evidence and arguments in depth; here I only summarize the main points.

3a. Formaldehyde Shows None of the Hallmarks Seen in All Known Leukemogens

Despite their disparate chemical nature, all recognized leukemogens cause pancytopenia, a decrease in all formed elements in the blood, at high doses (Goldstein 2009), but formaldehyde does not.

Similarly, in animal studies all leukemogens cause severe and overt bone marrow toxicity (Goldstein 2009, Golden *et al.* 2006), but formaldehyde does not.

There is no prototype of an agent that can cause leukemia without these other effects being seen as well, so there is no model for an alternative mode of action with any precedent.

3b. There are No Formaldehyde Adducts Outside of the Respiratory Tract

As noted, no evidence has been found that labeled formaldehyde caused detectable adducts outside of the immediate respiratory tract tissues. This contradicts the notion put forth by Zhang *et al.* (2008) that formaldehyde could be reactivated to a genotoxic form in bone marrow, and it confirms the lack of any significant systemic formaldehyde exposure from inhalation of the gas.

3c. There are No Chloromas in Humans, Which Would be a Consequence of Transformation of Precursor Myelopoietic Cells in Nasal Tissue

Zhang *et al.*, (2008) propose that precursor cells in nasal tissue might be affected directly by inhaled formaldehyde and then migrate to marrow, where they become finally transformed to leukemia cells. Goldstein, (2009) points out that, if precursor cells in nasal tissue were acted upon in this way, there should also be generation of chloromas, since thee isolated accumulations of myeloid tumor cells originate from the same proposed precursor cells in nasal tissue. There is no sign of chloromas among formaldehyde-exposed workers, however.

In sum, none of the proposed mechanisms by which inhaled formaldehyde could cause leukemia has any evidence in its favor, and each has considerable evidence against its operation. If formaldehyde were to be a human leukemogen, it would have to be so by some unprecedented mechanism that applies to no

other known leukemogenic agents. It would have to do so in a way that did not leave traces (in the form of adducts) of reactive formaldehyde in hematopoietic tissues or anywhere else other than the respiratory tract. If formaldehyde acts on hematopoietic precursor cells in the respiratory tract, it must do so in a way that fails to lead to other expected consequences of such actions, such as the generation of chloromas.

In conclusion, it does not seem possible to construct even a hypothetically plausible means for formaldehyde to cause hematopoietic cancers. If there is no means to do so, any tentative associations seen in epidemiological studies must be due to other causes, confounders, or chance.

4. Formaldehyde Inhalation Does Not Cause Leukemia in Animals

In no animal bioassay has inhaled formaldehyde, even at high doses that are markedly cytotoxic to the respiratory tract, caused leukemias or other hematopoietic cancers, or for that matter any systemic tumors. There are other chemical agents that do cause leukemias in rats, and all known human leukemogens are positive in animal studies.

The lack of effect in animals is despite the fact that rodents share the basic hematopoietic machinery with humans. It is not clear, then, why an effect in humans would not be paralleled in rodents. Formaldehyde is direct acting, so there is no question of differences in metabolic pathways or lack of local metabolic activation to explain the discrepancy.

Zhang *et al.* (2008) cite rat data to support their proposal that hematopoietic precursor cells exist in nasal tissue. If such cells are acted upon by formaldehyde in humans, why are they not also acted upon in rats, especially in view of the extraordinarily high and ongoing exposures experienced by rats at the higher bioassay doses? That is, by direct observation it is seen that the proposed mechanism whereby formaldehyde affects hematopoietic precursor cells in nasal tissue does not operate in rats. In view of this, it is not clear why it should be plausible to say it operates in humans.

It is important to underscore that what is at issue here is *not* simply that experimental animal data fail to provide added evidence of potential to cause hematopoietic cancers – that is, it is not just the inability to add "sufficient evidence in animals" or even "limited evidence in animals" to the checklist. It is that the negative animal evidence bears directly on the interpretation of the body of epidemiological evidence. If one is to use human evidence to override the biological plausibility problems and the lack of parallel

effects in animals, this corresponds to an assertion that there is some yet to be discovered, and yet to be plausibly hypothesized human-specific biological mechanism that allows for causal effects in humans where none are evident in animals or in mode of action.

5. In View of the Above, Formaldehyde Cannot be Judged to be a Potential Human Leukemogen

Epidemiology can at best show associations, and because of the vagaries and complexity of real-world circumstances, some inconsistency among studies is to be expected. The task in weighing the evidence regarding a possible causal role of an agent in disease appearance is to determine whether a causal interpretation is clearly preferable to alternative interpretations in which chance and confounding factors lead to apparent but non-causal associations.

Part of the evaluation concerns the body of epidemiologic evidence itself – whether the apparent effects are strong, consistent among studies, and compatible in pattern (regarding dose measures and dose-response gradients) with what would be expected from a causal agent. As has been shown, for formaldehyde the apparent effects are weak in some studies and absent in others. Effects are lesser in industries with higher exposures, the opposite of what one would expect of a truly causal agent. Meta-analyses tend to show that the collective effect across studies is null – there is no tendency toward the same direction that appears across studies despite the vagaries of each. Moreover, the finding of apparent associations seems to have depended on trying a large number of parallel analyses and picking out those that seem positive, skewing the interpretation of statistical significance. The dose measure apparently marginally associated with leukemias in the NCI study, "peak" exposure, is a hypothetical assignment of possible exposure to peaks rather than an actual measurement, while more biologically plausible measures of exposure, such as cumulative exposure, average concentration, and even the number of high peak exposures over life show no association.

In short, the epidemiologic data in themselves show little compelling evidence of a meaningful association of formaldehyde with leukemias or other hematopoietic cancers.

The larger aspect of the weight of evidence is to examine how the epidemiologic results jibe with the larger body of animal testing results, mechanism of action studies, and general knowledge about the properties of the agent and the biology of the endpoints in question. These data are not merely ancillary –

they are not separate pillars of a three-legged weight-of-evidence stool – they are (or ought to be) integrated into the interpretation of the epidemiological evidence itself. The epidemiological data *per se* can only show apparent associations; judging whether they are causal associations or better attributed to chance and confounding entails not only examining how strong the epidemiological results are in themselves, but also whether the causal mechanisms that it is necessary to hypothesize to judge the human data as supportive of causation stand up to the wider evaluation of biological plausibility.

The chemical properties of formaldehyde limit the plausibility of its potential action on hematpoietic processes. The toxicity of formaldehyde and the formation and maturation of blood elements take place in different tissues, and there is not a clear way for the one to influence the other. Hypothesized ways around this apparent refutation themselves do not stand up to the evidence, and the lack of any sign of a parallel process in animals, despite the presence of the same hematopoietic processes, shows that they do not operate.

These considerations lead to the scientific judgment that the occasional apparent associations of formaldehyde with leukemias or other hematopoietic cancers in some (but not most) studies, with some (but not most biologically plausible) dose measures but not with integrative measures of exposure, and with some (but not other) statistical models are much better interpreted as chance or spurious findings. If one is to take these results as evidence of causation, it is necessary to adopt as part of the weighing of evidence some reasoning by which the biological implausibility of a causal effect can be overcome and by which the lack of a comparable process in experimental animals comes about. That is, if one asserts causality despite these questions, the proposed scientific answer to those questions becomes a critical part of the weight of evidence, and the evidence for or against such explanations becomes as important as the epidemiological data themselves.

I conclude from such considerations that formaldehyde cannot plausibly be a human leukemogen, and that apparent associations in human studies are in fact much better understood as artifacts of analytical choices or chance findings. Accordingly, hematopoietic cancers do not add to the weight of evidence regarding formaldehyde as a potential human carcinogen.

6. The Evidence does Not Support Reclassifying Formaldehyde as a Known Human Carcinogen

The NTP criteria (NTP, 2009) for a "known human carcinogen" are:

> *There is sufficient evidence of carcinogenicity from studies in humans, which indicates a causal relationship between exposure to the agent, substance, or mixture, and human cancer.*

This is in contrast to the criteria for "reasonably anticipated to be human carcinogen," specifically:

> *There is limited evidence of carcinogenicity from studies in humans, which indicates that a causal interpretation is credible, but that alternative explanations, such as chance, bias, or confounding factors, could not adequately be excluded....*

The key to the difference between categories is the relative plausibility of causal and alternative explanations of the array of results in human studies. This is made explicit by Bradford Hill in his seminal paper on distinguishing causality from association:

> *None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on* ***the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?*** (Bradford Hill, 1965) [emphasis added]

That is, weight of evidence is not the simple counting up of apparently positive effects or the dismissal as unproven of alternative explanations. One has to examine an overall interpretation and tentative explanation of the array of results – positive, negative, and null – and one has to consider if the invoked causal explanation is a notably better way to account for all of these results than is offered by alternative explanations (which for epidemiology consists of confounding, chance, bias, other exposures or causes, and the inappropriate focus on the single analysis among many parallel analyses of the same data that appear to yield effects). Doing this entails not only examination of the human studies themselves, but also of the biological plausibility of the tentative explanations of how a causal effect could operate, judged by other experimental data and wider biological knowledge. All of these comprise the basis to decide whether the causal or the alternative explanations are more compelling.

That is, one needs to look at consistency, plausibility, and compatibility with wider knowledge to ask:

- Are the positive results compelling in themselves?

- Are they pulled from a sea of null results? (*i.e.*, has one shopped for a statistically significant response and then attributed significance to the choice after the fact?)
- What contrary observations (or lack of expected other outcomes) need to be explained in order to hold to the causal interpretation, and are such dismissals plausible? Are the causal explanations *ad hoc*, in the sense of being constructed to comport with the data and dodge apparent refutations by adding fillips or subsidiary hypotheses about causal processes?
- What is plausibility of other explanations of the array of outcomes? Is the causality by the chemical really compellingly better than such alternatives?

Seen in this way, the leukemias, as argued above, do not add to the weight of evidence. The human data themselves do not make a case for a credible association, and the mechanistic and animal data make clear that alternative explanations of the pattern of effects seen in human studies are far more credible than that formaldehyde is causing leukemias.

For the nasopharyngeal cancers, the case from the human data is at best limited, as has been judged before in the 11th ROC. Since that judgment, further meta-analysis has cast larger doubt on whether there is any credible association of nasopharyngeal cancer and formaldehyde exposure in humans. The apparent association is attributable to a single facility, with other facilities not showing effects, and the cancers seen in that facility are more plausibly attributed to metalworking exposures other than to formaldehyde. In short, the nasopharyngeal cancer case is weaker than it was in the 11th ROC.

Accordingly, there is no credible scientific basis to change the classification of formaldehyde to that of a known human carcinogen.

References

Agency for Toxic Substances and Disease Registry (ATSDR). 1999. *Toxicological Profile for Formaldehyde.* US Department of Health and Human Services, Public Health Service.

Bachand, A; Mundt, KA; Mundt, DJ; Montgomery, RR. 2010 (in press) Epidemiological Studies of Formaldehyde Exposure and Risk of Leukemia and Nasopharyngeal Cancer: A Meta-Analysis. *Critical Reviews in Toxicology.*

Beane Freeman, LE; Blair, A; Lubin, JH; Stewart, PA; Hayes, RB; Hoover, RN; Hauptmann, M. 2009. "Mortality from lymphohematopoietic malignancies among workers in formaldehyde industries: The National Cancer Institute Cohort." *J. Natl. Cancer Inst.* 101(10):751-761.

Blair, A; Zheng, T; Linos, P; Stewart, PA; Zhang, YW; Cantor, KP. 2001. Occupation and leukemia: a population-based case-control study in Iowa and Minnesota: *Am. J. Ind. Med.* 40(1):3-14

Bosetti, C; McLaughlin, JK; Tarone, RE; Pira, E; La Vecchia, C. 2008. "Formaldehyde and cancer risk: a quantitative review of cohort studies through 2006." *Annals of Oncology.* 19:29-43

Casanova, M; Heck, Hd'A; Everitt, JI. 1988. "Formaldehyde concentrations in the blood of Rhesus monkeys after inhalation exposure." *Food Chem. Toxicol.* 26:715-716.

Collins, JJ; Lineker, GA. 2004. "A review and meta-analysis of formaldehyde exposure and leukemia." *Regul. Toxicol. Pharmacol.* 40(2):81-91.

Golden, R; Pyatt, D; Shields, PG / Formaldehyde as a potential human leukemogen: An assessment of biological plausibility. / 2006

Goldstein, BD. Undated. "Clinical and hematotoxicological evaluation of current evidence does not support formaldehyde as a cause of human leukemia." *Hum. and Exp. Tox.* [Manuscript].

Hauptmann, M; Lubin, JH; Stewart, PA; Hayes, RB; Blair, A. 2003. "Mortality from lymphohematopoietic malignancies among workers in formaldehyde industries." *J. Natl. Cancer Inst.* 95(21):1615-1623.

Hauptmann, M; Lubin, JH; Stewart, PA; Hayes, RB; Blair, A. 2004. "Mortality from solid cancers among workers in formaldehyde industries." *Am. J. Epidemiol.* 159(12):1117-1130.

Heck, H; Casanova-Schmitz, M; Dodd, P; Schachter, E; Witek, T; Tosun, T. 1985. "Formaldehyde (CH2O) concentrations in the blood of humans and Fischer 344 rats exposed to CH2O under controlled conditions." *Am. Ind. Hyg. Assoc. J.* 46(1):1–3.

Hill, AB. 1965. "The environment and disease: Association or causation?" *Proc. R. Soc. Med.* 58:295-300.

International Agency for Research on Cancer (IARC). 2006. *Formaldehyde, 2-Butoxyethanol and 1-tert-Butoxypropan-2-ol.* IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. WHO, IARC, Lyon, France.

Kerns, WD; Pavkov, KL; Donofrio, DJ; Gralla, EJ; Swenberg, JA. 1983. "Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure." *Cancer Res.* 43(9):4382-4392.

McGregor, D; Bolt, H; Cogliano, V; Richter-Reichhelm, HB. 2006. "Formaldehyde and glutaraldehyde and nasal cytotoxicity: Case study within the context of the 2006 IPCS Human Framework for the Analysis of a cancer mode of action for humans." *Crit. Rev. Toxicol.* 36(10):821-35.

Marsh, GM; Youk, AO. 2005. "Reevaluation of mortality risks from nasopharyngeal cancer in the formaldehyde cohort study of the National Cancer Institute." *Regul. Toxicol. Pharmacol.* 42:275-283.

Marsh, GM; Youk, AO; Morfeld, P. 2007. "Mis-specified and non-robust mortality risk models for nasopharyngeal cancer in the National Cancer Institute formaldehyde worker cohort study." *Regul. Toxicol. Pharmacol.* 47(1):59-67.

Marsh, GM; Youk, AO; Buchanich, JM; Erdal, S; Esmen, NA. 2007. "Work in the metal industry and nasopharyngeal cancer mortality among formaldehyde-exposed workers." *Regul. Toxicol. Pharmacol.* 48(3):308-319.

Monticello, TM; Swenberg, JA; Gross, EA; Leininger, JR; Kimbell, JS; Seilkop, S; Starr, TB; Gibson, JE; Morgan, KT. 1996. "Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells." *Cancer Res.* 56(5):1012-22.

National Industrial Chemicals Notification and Assessment Scheme (NICNAS). 2006. "Formaldehyde: Priority existing chemical assessment report no. 28." A report of the Australian Government Department of Health and Aging. Last accessed at http://www.nicnas.gov.au on October 15, 2009.

National Toxicology Program (NTP). 2005. "Eleventh Report on Carcinogens (Table of Contents)." Accessed on April 8, 2005 at http://ntp-server.niehs.nih.gov/ntp/roc/toc11.html, 6p.

National Toxicology Program (NTP). 2009. "Report on Carcinogens: Background Document for Formaldehyde (Draft)." 593p., September 23.

Partanen, T; Kauppinen, T; Luukonen, R; Hakulinen, T; Pukkala, E. 1993. "Malignant lymphomas and leukemias, and exposures in the wood industry: An industry-based case-referent study." *Int. Arch. Occup. Environ Health.* 64(8):593-596.

Pyatt, D; Natelson, E; Golden, R. 2008. "Is inhalation exposure to formaldehyde a biologically plausible cause of lymphohematopoietic malignancies?" *Regul. Toxicol. Pharmacol.* 51:119-133.

Zhang, L; Steinmaus, C; Eastmond, DA; Xin, XK; Smith, MT. 2009. "Formaldehyde exposure and leukemia: A new meta-analysis and potential mechanisms." *Mutat. Res.* 681(2-3):150-168.